Increasing The Safety And Reliability Of Autonomous Vehicles

Yasin Erdal - MA, Digital Experience Design
Manchester, UK
The last decade has witnessed a huge growth in research and development of autonomous vehicles (AVs) and artificial intelligence (AI). However, the production of advanced technology has raised concerns over the risks and safety issues concerning AVs and AI which can consequently affect public perceptions of modern technology. The public have been found to be ill-informed about developments regarding AVs and AI, mainly only hearing about such developments through popularised news. Meanwhile, technology providers have often ignored a human-centered approach, rather than question people’s concerns and needs, have focussed solely on being first to deliver the service. Through the ‘design thinking’ approach, this paper aims to demonstrate a holistic solution for the issues which are closely related to each other by creating a binding experience. The findings focus on the development speed of autonomous vehicles’ decision making technology through AI. Image annotation technique is recognized as being the most important factor in the improvement of datasets which affect decision making performance in AVs. Current techniques to improve annotation speed and quality are therefore time-consuming and impractical. Findings suggest that advancements in both in-road infrastructures and AI can empower the autonomisation of vehicles, thorough research and safety implementations, and can increasingly change public perceptions to one of interest rather than one of fear. Current techniques offered to improve annotation speed and quality are time-consuming and impractical and this slows down the advancement process, costing millions of dollars. However, the concept of a digital platform, as discussed throughout this research, has the power to create a fun and educational experience for both AI and the general public. This training process encourages users to annotate visuals through gamification which has a significant potential to advance the technology remarkably faster and more effectively.
I want to take a moment to thank each one of the participants and supporters who contributed to this thesis by sharing their feedback, expertise and love. Their support and input were invaluable and allowed me to finalize this thesis.

Thank you;

My wonderful wife Rimma and my lovely parents Irina, Afyer and Adam for helping me to have this opportunity, Martijn Bout; ITRL Project Manager and my mentor, Jelle Disseldorp; My IRP body, Mia Kristensen, A.J. Huxtable-Lee and other Hypers, Judi Spiers and Tash Willcocks for their guidance and supervision, E. Jane Wardle; my beloved friend with super editing skills, Lydia Howland for her never-ending support and recommendations, Dmytro Smolovik for always being there and never complaining about me being a headache, Ruben Polderman; Project Manager at Municipality of Amsterdam, Jens Dahlstrøm Iversen; Innovation Consultant “Smart Cities at DTU”, Eetu Rutanen; Helsinki Metropolia University, Aldo Teran, Filip Klaessan; Roboticist, Frank Jiang, Umar Chughtai; Machine learning engineers, Ilkay T., Rimma T., Jani K., Elia F., Cristina E., Nicola L., Varun A.; participants, Samer Al Moubayed; CEO at Furhat Robotics, Anna Pernestål; Director at Integrated Transport Research Lab, and the entire team at the research lab for hosting me and providing constant support during my research project.
# CONTENTS

## 1. INTRODUCTION

1.1 Research questions ...................................................... 9  
1.2 Aim and objectives ...................................................... 10  
1.3 Methodology .......................................................... 11

## 2 LITERATURE REVIEW

2.1 Definition and recent history of AVs ............................................. 16  
2.2 Why would we need AVs? .................................................. 22  
2.3 AVs’ challenges and issues? ................................................ 26  
2.4 AI factor .......................................................... 27  
2.5 Technology enablers ...................................................... 32  
2.6 Public opinion ........................................................ 33  
2.7 Using design thinking to solve engineering problems ...................... 34  
2.8 Ethical issues ...................................................... 36

## 3 DISCOVERY

3.1 Research ........................................................ 38  
3.2 Opportunity and ideas .................................................. 46  
3.3 Understanding the AI .................................................. 47  
3.4 Problem definition .................................................... 50  
3.5 Ideation .......................................................... 51  
3.6 Final idea .......................................................... 56

## 4 ETHICS

59

## 5 CONCLUSION

60

## 6 APPENDICES

73
List of Abbreviations:

ACC  Adaptive Cruise Control
ADAS  Advanced Driver Assistance Systems
ADS  Automated Driving System
AI  Artificial Intelligence
A/IS  Autonomous and Intelligent Systems
AV  Autonomous Vehicle
AVs  Autonomous Vehicles
CAV  Connected and Automated vehicle
C-V2X  Cellular Vehicle to Everything
GPS  Global Positioning System
IOT  Internet of things
ITS  Intelligent Transport System
ITRL  Integrated Transport Research Lab
I2V  Infrastructure to Vehicle
NHTSA  National Highway Traffic Safety Administration
SDV  Self-Driving Vehicle
SDC  Self-Driving Car
SDS  Self-Driving System
S.M.A.R.T.  Specific, Measurable, Attainable, Relevant and Timely
V2I  Vehicle to Infrastructure

Glossary of Terms:

Big Data  extremely large data sets that may be analysed computationally to reveal patterns, trends, and associations, especially relating to human behaviour and interactions.

Internet of Things  The Internet of things is the network of devices, vehicles, and home appliances that contain electronics, software, actuators, and connectivity which allows these things to connect, interact and exchange data.

Machine learning  A subset of artificial intelligence
Ever-evolving communication, partnership, and investment among companies aim to boost the development speed of emerging technologies on a global level (CB insights, 2018) and vehicles with a different level of automation have become part of these emerging and disruptive Technologies by increasingly participating in public traffic. The continuing evolution of automotive technology aims to deliver even greater safety benefits and Automated Driving Systems (ADS) that one day can have the ability to drive us anywhere we want when we do not wish or cannot drive it ourselves (NHTSA, 2018). Some companies have already covered hundreds of thousands of miles testing their vehicles on public roads while others have been using short range or test tracks to observe and develop the technology. Governments and research institutes support or take part in the development of autonomous technology as Connected and Automated Vehicles (CAV) are the key factors in paving the way for smarter cities with a reduced carbon footprint and bigger spaces for recreational activities (Open Access Government, 2018). The success of Autonomous Vehicles (AVs) in changing the roads is inevitable and that alone raises some critical questions. Many people are concerned about the safety implications and the consequences of Autonomous Vehicle (AV) technology, especially as automotive companies and AV startups often focus on delivering technology first, thus subsequently neglecting the human factor of the equation. As the U.S. Secretary of Transportation, Elaine Chao highlights, “while the technology holds promise, it has not yet won public acceptance” (The Verge, 2018). Even though the public is one of the major factors to truly understand and serve, a holistic approach to understand each stakeholder’s concerns and find commonalities is also very important. As Freeman (2010) states, “It is this bonding or binding idea that is most interesting to explore”, a holistic observation and approach towards the problems of stakeholders as “bound together by the jointness of their interests”. Each stakeholder, as well as different factors of technology and safety, are examined and will be discussed in this report. The role and potential impact of ‘design thinking’ in understanding and solving problems of different stakeholders will also be outlined.
“Design thinking can be described as a discipline that uses the designer’s sensibility and methods to match people’s needs with what is technologically feasible and what a viable business strategy can convert into customer value and market opportunity.”

— Tim Brown CEO, IDEO
Knowing that autonomous technology might have the potential to reduce traffic accidents by nearly 90% (Open Access Government, 2018), I will discuss the hindering factors and explore whether people would still rely on an intelligent system to drive their cars at an average speed. As this emerging technology is being tested and gradually allowed on the roads under controlled conditions, the focus will be on practical technological solutions and their social consequences (Holstein, Crnkovic and Pelliccione, 2018). Existing issues and the opportunity areas will be examined to increase safety and create a binding experience among the stakeholders to enable the delivery of AV technology sooner and safer.
1.1 Research questions

How might we increase the safety and reliability of AVs?

How might we design a binding experience among AV stakeholders?

Why?

Transitioning into a whole new technology which will impact people’s lives is challenging. When it comes to autonomous technology, people have different thoughts, needs and concerns and it is not only consumers that display these concerns. Technology providers, policy makers, investors, and researchers, all have different types of questions in mind and answering some of these questions helps to alleviate some of the concerns and gain more acceptance. Otherwise, consequences might delay the approval of the technology even though the environmental and safety benefits of AVs are promising (Smith and Vardhan, 2018, p.13). The outcome or the possible effects of AVs on human lives are not so easy to predict, however, working on this cause of technological development can be beneficial as AVs have a great potential for saving millions of lives based on the fact 94% of serious crashes are due to human error (NHTSA, 2018). Designing a binding experience can therefore, inform people better, smooth the process and promote the adoption of AVs across the world by global consumers.
Aim

The primary aim of this research is to increase the safety and public acceptance of AV technology by presenting an innovative solution to the stakeholders. Additionally, to present solid evidence on how design thinking can play a major role in solving complex and technical problems.

Objectives

Using design research methods to observe and understand the different factors which have a negative influence on technological development. Designing a feasible, viable, and desirable solution offers to ease the adoption process of AVs.

Drawing attention to infrastructure-based solutions to increase safety by transmitting relevant information to AVs in alternative ways.

Designing a reliable platform and a binding experience through the co-existence of Artificial Intelligence (AI) and humans to increase annotation quality and speed in machine learning.
Throughout my research project, I conducted thirteen in-depth interviews face to face in five countries involving most of the stakeholders, plus one interview that was conducted through video chat as the participant was unable to travel for a face-to-face interview. I also attended a conference specifically aimed at AVs, had a test ride in three different autonomous vehicles and observed people riding in them. In addition to this, I read over one hundred journals, articles, conference proceedings, reports and books in order to gain a general understanding of AV landscape and current technologies. I used the ‘double diamond’ design process model; a visual map to organize, digest and process all the data and information gathered as input (Figure 1).

Design thinking and design research methodologies were therefore integrated throughout the whole research. As a design researcher I leveraged participants’ information in generating a new project definition to identify new goals. In order to do this, it was essential to understand the participants and their experience holistically by observing psychological, physiological, cultural and social forces to provide me with useful insights (Chayutsahakij, 2018). Bell (2009), former director of the User Experience Group within Intel Corporation, suggests that to be successful in a contextual research, the approach has to be relevant, and to be relevant, the environment, conditions, culture, and the habits of people must be considered, so we can see the world through their eyes and therefore understand them better. I narrowed down my research area and reframed the problems several times following the contextual interviews. I also tried a new method which I had never heard of to unlock what I was unable to see in the final phase of my research and interviewed real AI models including a real robot, Google assistant, several chatbots, and two AI engineers who were asked to think and act as an advanced AI and answer questions accordingly.

1.3 Methodology

Figure 1. Double diamond design process model

Design Council, 2007
“Great research breeds great design. Design researchers dream up new ways to spark and distill insight.” IDEO
During the four months of research, autonomous and intelligent systems (A/IS), as well as its effects on individuals and society, were investigated and evaluated. The findings and opportunity areas were discussed, evaluated and validated by the engineers and researchers at the Integrated Transport Research Lab (ITRL) at the Royal Institute of Technology (KTH). To ensure A/IS research outcomes are aligned to be beneficial for society, research and design is underpinned by ethical and human-centered values. Value-based system design approach which puts human advancement at the core of A/IS development was adopted during the entire project term with such methods recognizing that machines should serve humans and not the other way around (Ethically Aligned Design, 2017). IDEO’s “sacrificial concepts” method was used to evaluate my ideas by designing four diverse concepts. The concepts were tested by the users and validated by the experts to prevent time loss before creating a prototype. I abandoned some of my ideas due to the participant and expert feedback and decided to embrace the ones which were found more appealing and feasible for both users and experts. Feedback was taken into account and following the iteration of this process, I focused on one final concept to develop. The process and possible solutions were developed based on the experience design principles where the digital product was only one touch point and part of a holistic experience. Experience design leader, Schwartz (2017, p.12) states “Experience design is concerned with developing a holistic understanding of the relationship between a person and product over time, meeting needs and exceeding expectations in ways which users perceive as valuable, effortless, and emotionally satisfying.” For the best possible outcome, the research and the solution were AI and human-centered since the coexistence of these two concepts appeared more beneficial based on the evidence (Koizumi, 2016).
In this report, the terms ‘autonomous vehicle’, ‘self-driving vehicle’, ‘self-driving car’ and ‘driverless car’ were used interchangeably. Apart from the experts, people who were involved in the interviews, product test, feedback and/or validation are referred as a participant.

Despite having an academic background in computer technology and programming, AVs and AI is not my area of expertise. It was, therefore, difficult to find an expert to support this research and I had no possibility to validate my ideas without specific supervision in the first two months of the project. The first two months were spent intensively studying these technologies and conducting primary research. Luckily a collaboration with experts at the Integrated Transport Research Lab, Royal Institute of Technology was organised two months prior to the project delivery in Stockholm, Sweden. Due to time constraints, Specific, Measurable, Attainable, Relevant and Timely (S.M.A.R.T) criteria was applied throughout the project as S.M.A.R.T goals help to clarify ideas, focus efforts and enable the use of time and resources productively in order to achieve the full potential (Mindtools, 2016).
There is a vast amount of literature on AVs especially in recent years which has seen a growing interest in different aspects of autonomous technology. This mostly covers AI, ethics, feasibility, infrastructure, allocation and CAV running through cellular vehicle to everything (C-V2X) which is powered by fifth generation (5G) of cellular broadband mobile communication and Internet of Things (IoT). However, not many studies have been published that draw attention to the importance of the design thinking approach in AV research. Formerly Google now Waymo’s AVs had already driven 10 million miles on public roads in the United States by October 2018 (Waymo, 2018). There is an aggressive race to reach the fully autonomous level among the automotive and technology companies but they often ignore human factors and values during the process. As A/IS technologies are being implemented in an increasing number of existing vehicles, it is necessary to include people and address their concerns and issues in this process before it is too late (Cunningham and Regan, 2015).
Autonomous, automated, self-driving or driverless terms often refer to a kind of intelligent car which drives to a specific destination based on the information obtained from its sensors and cameras. This includes the perception of the path environment and information of the route and car control. The main characteristic of an AV is transporting people or objects to a predetermined target without human interaction (Zhao, Liang and Chen, 2018). However, the autonomous term can also refer to various vehicles other than cars. Trucks, drones, busses, aeroplanes, trams, trains or ships are some of these vehicles that can be or will be operated autonomously without any human intervention. Fully automated vehicles that drive us instead of us driving them will be on the roads in the very near future. The SDV will be able to integrate onto U.S. roadways by progressing through six levels of driver assisted technological advancements in the coming years. The international standard for automation levels which were defined by Society of Automotive Engineers (SAE) was also adopted by National Highway Traffic Safety Administration (NHTSA). The automation levels (following page) include everything from non-autonomous where a fully engaged driver is required at all times, to full autonomy where an AV operates independently without a human driver (NHTSA, 2018).
# Levels of automation

<table>
<thead>
<tr>
<th>Levels of automation</th>
<th>Who does what and when?</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Level 0</strong></td>
<td>The human driver does all the driving (non-autonomous).</td>
</tr>
<tr>
<td><strong>Level 1</strong></td>
<td>An advanced driver assistance system (ADAS) on the vehicle can sometimes assist the human driver with either steering or braking/accelerating, but not both simultaneously.</td>
</tr>
<tr>
<td><strong>Level 2</strong></td>
<td>An advanced driver assistance system (ADAS) on the vehicle can control both steering and braking/accelerating simultaneously under some circumstances. The human driver must continue to pay full attention (&quot;monitor the driving environment&quot;) at all times and perform the rest of the driving task.</td>
</tr>
<tr>
<td><strong>Level 3</strong></td>
<td>An Automated Driving System (ADS) on the vehicle can perform all aspects of the driving task autonomously under some circumstances. In those circumstances, the human driver must be ready to take back control at any time when the ADS requests the human driver to do so. In all other circumstances, the human driver performs the driving task.</td>
</tr>
<tr>
<td><strong>Level 4</strong></td>
<td>An Automated Driving System (ADS) on the vehicle can itself perform all driving tasks and monitor the driving environment – essentially, do all the driving – in certain circumstances. The human need not pay attention in those circumstances.</td>
</tr>
<tr>
<td><strong>Level 5</strong></td>
<td>An Automated Driving System (ADS) on the vehicle is fully autonomous and can do all the driving in all circumstances. The human occupants are just passengers and need never be involved in driving.</td>
</tr>
</tbody>
</table>
Today’s AVs are not made of a single technology, but rather, involves a complex system that integrates multiple technologies (Figure, 2). These technologies consist of three major subsystems of algorithms; sensing, perception and decision making.

The algorithm subsystem extracts meaningful information from sensors’ and camera’s raw data to understand its environment and make decisions about its actions (Tang et al., 2018).
During the 1980s and 1990s, pioneering researchers around the world including Tsukuba Mechanical Engineering Lab (Japan), Ernst Dickmanns (Germany), and Navlab: The Carnegie Mellon University Navigation Laboratory (U.S.) built the fundamentals of today’s autonomous technology (Kurban and Lipson, 2016). The vision-guided driverless Mercedes-Benz robotic van, which was designed by Ernst Dickmanns (Figure 4) and his team at the Bundeswehr University Munich, Germany, achieved a speed of 63 km/h on streets without traffic. The newer autonomous vehicles have become more and more efficient with time. The twin robot vehicles VaMP (Figure 3) and Vita-2 of Daimler-Benz and Ernst Dickmanns, in 1991 drove more than 1,000 km on a Paris three-lane highway in standard heavy traffic at speeds up to 130 km/h, but semi-autonomously with human interventions. They demonstrated various functions of autonomous driving such as in free lanes, convoy driving, and lane changes with autonomous passing of other cars (Bimbraw, 2015).

The Defense Research Advanced Projects Agency (DARPA) announced a Grand Challenge in the United States in 2004, 2005, and 2007, which provided a lead prize of $1 million for whoever could build the fastest SDC that could drive 142 miles through the Mojave Desert (WIRED, 2018). This project brought teams together from around the world in order to succeed in the production of the required vehicle and this collaboration of multiple agencies such as government, universities and car manufacturers, ranging from Japan to Europe, resulted in more progressive outcomes for technological advancement (Lari, Douma and Onyiah, 2015). Those who participated in these challenges still form some of the core researchers and engineers and some of them moved on to work on AV technology at Google, Uber, Tesla Motors and a host of startups seeking to make SDVs a reality (The Economist, 2018). Google’s SDC project which started in 2009 was the first signal of a commercialized autonomous race (Waymo, 2018) and Tesla Motors became the first winner by releasing a software update to its Model S vehicle, enabling a feature dubbed “Autopilot” in October 2015. The software enabled the car to autonomously follow a lane, as well as switch lanes or park on command and these abilities placed the Model S among the first publicly available “SDC” (Engineering, 2018).
Most of the automotive companies invested hundreds of millions of dollars to be able to release level 4 and level 5 AVs in 2020 and 2021 (TechEmergence, 2018). Innovations generally follow a predictable deployment pattern, often called an S-curve or Gal’s Insight, as illustrated in figure 5. They generally start with a concept which requires testing and approval, experiences of an initial commercial release, product improvement, market expansion, differentiation, maturation, and saturation. AV technology is expected to follow the same pattern with a rapid transition (Litman, 2018).

![Innovation S-curve](Litman, 2018, p. 18)
Experience with mass motorization and previous vehicle technology deployment can be used to model autonomous vehicle implementation. Level 4-5 AVs will become commercially available in the 2020s but are initially limited in performance and are expensive. Due to these limitations, autonomous vehicle sales are unlikely to be of immediate high demand, however, with market shares increasing in subsequent decades as their performance improves, as well as declining prices and improved consumer confidence, this demand will begin to rise. By the 2040s approximately half of the vehicles sold and 40% of vehicle travel could be autonomous. Without mandates, deployment might be similar to automatic transmissions, and a considerable number of drivers may continue to choose manual transmissions due to personal preferences (Litman, 2018). On the other hand, experts of the Institute of Electrical and Electronics Engineers (IEEE) estimated that up to 75% of all vehicles will be autonomous by 2040 while the AVs will be the most viable form of intelligent transportation (sites.ieee.org, 2018). AVs will save at least $60 billion to the American economy by preventing accidents and reducing traffic congestion (Duarte and Ratti, 2018). However, the implications of the technology may cause millions of drivers to lose their jobs. According to the American Trucking Association, there are currently 3.5 million professional truck drivers in America (Entrepreneur, 2018). The number of employees at risk exceeds 5 million including the other professional driving services according to research conducted by Harvard University (NBC News, 2018). Canada can also be considered as another example who have raised questions about the negative implications of AV technology. While the effect of AV technology on the Canadian economy is not yet fully clear, it could lead to job losses in sectors or occupations that employ over 1.1 million Canadians. Direct employment displacement would potentially affect industries such as: Truck and courier service drivers, taxi drivers, bus drivers, snowplough drivers, traffic police, traffic wardens, driving instructors, tow truck drivers as well as autobody repair mechanics, auto insurance agents and salespeople, parking attendants, and gas station employees (Chong and Sweeney, 2018).
2.2 Why would we need AVs?

Safety predictions

The Global status report on road safety 2015 reflects information from 180 countries and highlights that worldwide, the total number of road traffic deaths has plateaued at 1.25 million per year (World Health Organization, 2018) with 94% of serious crashes being due to human errors. Distracted driving such as texting while driving is the most common cause of road accidents in the U.S., resulting in more crashes every year than speeding, drunk driving, and other major accident causes (Strong Tie Insurance, 2018). Crashes in which at least one driver was identified as being distracted resulted in 3,267 fatalities, 735,000 nonfatal injuries and damaged 3.3 million vehicles in property-damage-only crashes in 2010. This represents about 10% of all motor vehicle fatalities and 18% of all nonfatal crashes and these crashes cost the U.S $39.7 billion in 2010 (Blincoe et al., 2010). AVs have the potential to remove human error from the crash equation, which will help protect drivers and passengers, as well as bicyclists, pedestrians and animals (NHTSA, 2018). A statistic from the Humane Society of the U.S. (HSUS) makes the claim that millions of wild animals are struck down by vehicles each week in the U.S. (The Dodo, 2018).
“As transportation is digitized in the next decade, driverless cars, guided by GPS and radar, will share our highways. ‘Traffic accidents’ and ‘traffic jams’ will become archaic terms. Thousands of lives will be saved every year.”

– Michio Kaku, Physicist
AVs could deliver additional economic and societal benefits. According to NHTSA, motor vehicle crashes in 2010 cost the U.S. $242 billion in economic activity, including $57.6 billion in lost workplace productivity, $594 billion to loss of life and decreased quality of life due to injuries and therefore by eliminating the vast majority of motor vehicle crashes, costs will also be significantly reduced. Under various levels of AVs, there is a potential for congestion savings due to adaptive cruise control (ACC) measures and traffic monitoring systems could smooth traffic flows by seeking to minimize accelerations and braking in freeway traffic. This could increase fuel economy and congested traffic speeds by 23–39% and 8–13%, respectively, for all vehicles in the freeway travel stream (Fagnanta and Kockelman, 2015). Moreover, Americans currently spend more than 6.9 billion hours a year sitting in traffic, according to the American Society of Civil Engineers and AVs could therefore free as much as 50 minutes a day for passengers, who will be able to spend the reduced travel time working, relaxing, or accessing entertainment. The time saved by commuters every day might add up globally to one billion hours (McKinsey, 2015).
Environmental effects

The reduction in congestion will also have environmental benefits as it will possibly result in a reduction of CO2 emissions as well. The AV software can be programmed to reduce emissions to the maximum extent possible which can be expected to contribute to a 60% fall in emissions (ITS digest, 2018). A great example of this innovative technology and improved fuel efficiency can be found in truck platooning. According to European Automobile Manufacturers Association (ACEA), the definition of truck platooning is the linking of two or more trucks in convoy, using connectivity technology and automated driving support systems. These vehicles automatically maintain a set, close distance between each other when they are connected for certain parts of a journey, for instance on motorways while mutually communicating through smart technology. This process requires less fuel consumption and therefore reduces CO2 emissions considerably.

Given that trucks can drive closer together, the air-drag friction is reduced significantly. Platooning can reduce CO2 emissions by up to 16% from the trailing vehicles and by up to 8% from the lead vehicle (Platooning roadmap, 2017). Since AVs will be able to drive themselves, this will eliminate the parking need in city centers and create more space for other activities (KTH Royal Institute of Technology, 2016). AVs also provide effective transportation options which can be very reasonable for people. Such options include facilitated mobility for a significant amount of the population; in particular, the elderly and people with physical disabilities who would be able to enjoy the freedom of travelling independently (BCG, 2018).
Even though AVs are already in production and are being tested on the roads, there are still many challenges such as those regarding the legal framework, regulatory changes, cost of technology, safety and trust, security and environmental factors (Howard and Dai, 2013). Many of these challenges are directly related to the technological advancement of the AVs’ software and hardware which can be solved by improving AV technologies. For example, lidar is a critical sensor for AVs. It is a sensing technology similar to a radar that detects objects with pulses of laser light even when the objects are not visible to the human eye. Today, almost all the AV manufacturers apart from Tesla motors, rely on the lidar technology which can cost thousands of dollars. This is why the size, complexity, and cost of the current generation of lidar sensors are significant obstacle which make AVs far from being affordable for the average consumer to buy (IEEE Spectrum, 2016). By advancing the Lidar technology, the size, weight, and the cost can be dramatically reduced. In 2017, Quanergy was the first company to advance the technology by using some of the same materials and manufacturing processes as computer chips. This enabled the company to develop relatively cheap Lidar sensors which were small enough to be embedded in car headlights and cost only $250 dollars (MIT Technology Review, 2017).

While the technology improvement for the hardware of the AVs is crucial in making it more cost effective and enable a better object detection capability, the software factor is even more important and critical to making AVs safe and reliable for humans. On March 18, 2018, an Uber test vehicle, operating with a self-driving system in computer control mode, struck a pedestrian, in Tempe, Arizona. According to data obtained from the Self-Driving System (SDS), the system first registered radar and LIDAR observations of the pedestrian about 6 seconds before impact, when the vehicle was travelling at 43 mph. As the vehicle and pedestrian paths converged, the SDS classified the pedestrian as an unknown object, as a vehicle, and then as a bicycle with varying expectations of an alternative travel path. At 1.3 seconds before impact, the SDS determined that an emergency braking manoeuvre was needed to mitigate a collision. Uber disabled the emergency braking manoeuvres while the vehicle was under computer control, to reduce the potential for erratic vehicle behaviour. The vehicle speed at impact was 39 mph (National Transportation Safety Board, 2018). The accident report is solid proof that Uber’s SDS had a very poor decision-making algorithm which made its AV a potential threat to human life back in March 2018. The accident could have been prevented since the object detection worked six seconds prior to the crash by informing the system. However, this brings up other questions regarding ethical issues. Studies may be aimed at enhancing our understanding of how to safely and efficiently re-engage the driver and in designing the automated vehicle from a human-centred perspective as this can be life-saving. Investigations also need to identify potential misuses of automation and develop countermeasures to reduce behaviours such as overreliance (Cunningham and Regan, 2015). Whether AVs improve public health, decrease traffic congestion and reduce climate change will depend on informed, science-based policy. This will require a robust research agenda and accessible data on the performance and operation of self-driving vehicles. The policy must, therefore, facilitate open data-sharing while ensuring that appropriate privacy protections are in place (Maximizing the Benefits of Self-Driving Vehicles, 2017).
Artificial intelligence is the most critical component which makes an AV operate safely on its own. If all the sensors of the AVs are meant to be the eyes of the vehicle, then the AI can be considered as the brain of the whole mechanism delivering the driving instructions. Regardless of the quality and performance of the hardware, AVs can be vulnerable and lethal if the AI in the vehicle has a poor algorithm, as outlined in the factual description of the Uber case above. In order to understand the limits, problems, and potentials of an AV, it is necessary to understand the way AI works and explore the technical challenges.

2.4 AI factor

Definition of AI

AI currently covers a huge variety of subfields ranging from general to specific tasks, such as playing chess, proving mathematical theorems, writing poetry, driving a car on a crowded street, and diagnosing diseases (Russell and Norvig, 2010). Feigenbaum and Barr (1981) define AI as, “the part of computer science concerned with designing intelligent computer systems, that is, systems that exhibit characteristics we associate with intelligence in human behaviour – understanding language, learning, reasoning, solving problems, and so on”. However there are a set of possible answers to the “What is AI?” question and the definition of AI can differ in terms of its goals (Stanford Encyclopedia of Philosophy, 2018). Basically, its interrelated goals can be discussed in the areas of science and engineering which cover older disciplines such as philosophy, logic/mathematics, computation, psychology/cognitive science, biology/neuroscience, and evolution, all which are rooted to AI. Scientific goals can determine which ideas about knowledge representation, learning, rule systems, search, and so on, and explain various sorts of real intelligence whereas an engineering goal can cover and solve real-world problems using AI techniques such as knowledge representation, natural language processing, automated reasoning, machine learning, computer vision, and robotics (Bullinaria, 2005).
Among the subsets of AI, machine learning is a modern technique for making predictions by mining information from big data. This means being capable of delivering faster and more accurate results by analyzing extremely large and complex data. It has been proven as one of the most successful achievements in AI with notable examples as in AVs and the famous AlphaGo, a computer program which surpasses the top human players at the strategy board game “Go” (Gao et al., 2018). Machine learning algorithms are extensively used in AVs and capable of learning through diverse data to identify patterns and make decisions with minimal human intervention (Sas, 2016). The current and future applications of machine learning are going to control AVs as well as evaluate the driver’s condition through data fusion from multiple sensors like lidar, radar, camera or the IoT (Internet of Things) devices. This allows an AV to have the capability of directing itself to a hospital if it notices that the driver’s condition is not good (Kdnuggets, 2017).
Another subset of the AI which is used in AVs is computer vision. Computer vision can be considered as the eyes of the AVs which perceive both external and internal objects. Szeliski (2010, p3) states "computer vision tries to describe the world we see in one or multiple images and to interpret its properties, such as shape, illumination, and colour distributions. It is amazing humans and animals can do this so effortlessly, while computer vision algorithms are error-prone". Computer vision has many uses, for example, the use of digital cameras to capture people’s smiles or the facial recognition on smartphones which allows the owner to unlock their phone. In AVs, computer vision can identify and distinguish objects on and around the road such as traffic lights, pedestrians, traffic signs, obstacles (Techopedia, 2016). Advanced computer vision algorithms contain a definition of objects. An algorithm looking for street signs will have some understanding of what a street sign is: its general shape, size, colours, and is programmed to recognize and deal with this information. The algorithm can use this information to better identify objects by matching outlines and patterns to a known set of images (Tufts, 2017). Humans learn through life experiences and can classify a cat from observing many cats throughout one’s lifetime. Differentiations can also be made between what constitutes spam email, for example, and therefore humans have the intellectual ability to identify different characteristics of an object based on observations and experiences. To help an AV to understand the world it is surrounded by, we need to give it enough sets of training data with the desired output. In machine learning, this approach refers to “supervised learning” (Data Science, 2017). If we want AI to be able to classify a cat, we need to train it with a diverse dataset with thousands of photos of cats. Cats with different colours, sizes, angles, yawning, jumping, stretching, running, and other identifiable features which can be programmed into AI in order to recognise it as a ‘cat’. This way, the classification algorithm of the AI learns a model of how the world of cats work (The New Stack, 2018).

In the last few years, the field of machine learning has made tremendous progress by achieving reasonable performance on hard visual recognition tasks, and even matching or exceeding human performance in some domains (TensorFlow, 2018). AVs technology is data driven where diverse, large-scale annotated visual datasets are the most vital elements for supervised learning tasks in computer vision. These datasets define the limits of an AV’s accuracy and performance. Typical deep learning models can require millions of training images to reach a high performance and annotating such massive amounts of data is one of the biggest technical challenges in today’s AV industry (Yu et al., 2018). It is, therefore, a critical task to properly prepare the datasets. There are many labelling tools and different options to build new datasets or improve the existing ones and to assure the quality, which could be done through in-house teams or professional data training companies. In both options, annotation of one object can take nearly 30 seconds, and creating a dataset with a million of objects could take around 347 days of work without sleeping or taking a break (Altexsoft, 2018). Preparing and improving the datasets takes nearly 80% of data scientists’ working time and most of them take it as the least enjoyable part of their job (Wilson, 2017).
Insurance Institute for Highway Safety (IIHS) tested five of the leading brands’ systems on track tests over hills and curves. According to the results (Figure, 7,8), while none of the AVs crashed, almost all of them missed the mark multiple times by crossing or touching lane lines or disengaging during driving (Quartz, 2018).

While the results are promising, it is also possible to say that the technology is still not ready and different technology providers have different success rates. Machine learning algorithms and datasets require improvement to prevent misidentification of the objects (Eykholt et al., 2018).
“I think people underestimate how much basic science still needs to be done before these cars or such systems will be able to anticipate the kinds of unusual, dangerous situations that can happen on the road,”

– Yoshua Bengio - AI expert
Autonomous vehicles are unable to interact with one another unless they use the same networking and communication infrastructure which permits the exchange of data, without this connection, AVs exclusively observe and react to their own environment. Using the more recent technologies will enable the interconnection between the vehicles and devices with a network infrastructure. Shared information can be used to prevent collisions with early warnings and facilitate a greater efficiency. Connected automated vehicles (CAV) and infrastructure will be in constant feedback circulation in real time and therefore the vehicles will be aware of congestion problems and dangerous zones to determine optimal routes and allocate themselves in the best possible way (WSP Group, 2016). CAV basically uses two different technologies; IoT and C-V2X (Cellular vehicle to everything). From the perspective of information technology, IoT is a huge global information system composed of hundreds of millions of devices that can be identified, sensed and processed based on various communication protocols. With the support of 5G mobile communication, IoT devices can autonomously connect, communicate, and interact with each other which provide management and control for decision-making, without any human intervention (Zhu et al., 2005).

On the other hand, C-V2X is a technology which enables the communication between the vehicles, objects and pedestrians. The technology is simply designed to send and receive critical traffic information through Vehicle-to-Vehicle (V2V), Vehicle-to-Infrastructure (V2I), Vehicle-to-Pedestrian (V2P) and Vehicle-to-Network (V2N) communication which uses the traditional cellular links to enable cloud services (Papathanassiou and Khoryaev, 2017). AVs supported by C-V2X communications have a great potential to save countless lives by preventing millions of traffic accidents. C-V2X technology will enable services such as hazard warnings, cooperative adaptive cruise control, platooning and more. Real-time communication informs vehicles of the invisible surroundings that neither the driver nor the vehicle’s built-in sensors can predict, enabling safer and more efficient driving (Nokia, 2017). C-V2X technology can be used in all vehicles with or without ADAS and ADS. Advanced communication and cooperation between cars and surrounding objects will be a part of everyday life and this will even reshape the roads in the future as the traditional physical existence of the traffic signal lights will not be needed any longer since all the vehicles will inform each other of their actions (Tettamanti, Varga and Szalay, 2016). However, in present time, it will not be profitable enough unless the majority of the automotive companies and vehicles in traffic switch to the C-V2X technology (Boban et al., 2017).
Even though AVs offer many benefits, the existing issues may raise new concerns amongst the public. While some studies, news articles, and automotive companies often present AVs as the ‘future of mobility’, there are also studies which address unavoidable concerns among car drivers towards this radical innovation (König and Neumayr, 2017). According to various studies, once an individual questions their trust in a technology, they are more likely to avoid this technology, reducing potential benefits and causing limitations in future use (Reimer, 2014). A key finding with much of the literature on public opinion is the data security concern. Results of an international questionnaire among 5000 respondents indicate numerous concerns mainly regarding software hacking/misuse followed by the safety, legal and data transmitting issues of AVs (Kyriakidis, Happee and de Winter, 2015). Public opinion needs to be taken into account by automotive companies and policymakers. When it comes to AVs, security has a paramount importance, and software security is a fundamental requirement. AVs will be able to connect each other as well as to any other device which is connected to an internet or cloud system. Interconnection of AVs and other devices will increase the safety and efficiency but it can actually make AVs even more vulnerable to cyber attacks. This risk is something that is yet to be explored and the implications of such risks are not yet known by technology providers and policymakers. In August 2017, the UK’s Department for Transport published a document “Key principles of vehicle cybersecurity for connected and automated vehicles”. It is built on eight basic principles:

1. Organizational security is owned, governed, and promoted at board level;
2. Security risks are assessed and managed appropriately and proportionately, including those specific to the supply chain;
3. Organizations need product aftercare and incident response to ensure systems are secure over their lifetime;
4. All organizations, including subcontractors, suppliers, and potential third parties, work together to enhance the security of the system;
5. Systems are designed using a defence-in-depth approach;
6. The security of the software is managed throughout its lifetime;
7. The storage and transmission of data is secure and can be controlled;
8. The system is designed to be resilient to attacks and respond appropriately when its defences or sensors fail (Holstein, Crnkovic and Pelliccione, 2018).

Today, what we know about public opinions regarding AVs is largely based on quantitative research which relies on structural data collection and consists of experiments that do not take place in natural settings. In addition, this method does not allow participants to explain their feelings, choices or the meaning that the questions may have for those participants (Simply Psychology, 2017). Having more data is therefore not helping us make better decisions on the existing issues. It is important to observe and understand emergent human dynamics that haven’t happened yet (TED, 2018). These dynamics and the future states of the complex AV problems can be observed and identified through design thinking (McKinsey & Company, 2016).
Razzouk and Shute (2012) define design thinking as an analytic and creative process that engages a person in opportunities to experiment, explore the issues, create and prototype models (Figure, 9), gather feedback, and redesign. Design thinking has no borders, on the contrary, it is flexible and all about keeping the bigger picture in mind while focusing on specifics. When looking at the different kind of innovation concepts, the most useful ones can be gathered under design thinking where creativity complements the scientific way of doing things (Owen, 2006).
Design thinking methods and tools can be implemented through the guidance of design process models with the ‘Double Diamond’ being one of these process models developed by an in-house research team at the Design Council in 2005. Double diamond has four distinct phases which were identified as discovery, problem definition, development and delivery. The divergent and convergent phases of each diamonds cover the entire journey of an innovation and product development process by guiding the design thinker through each of these stages (Design Council, 2015). Following the design process simplifies the journey and offers diverse ideas based on the root of problems and needs. As a result these ideas can transform into meaningful products and services. The impact of design thinking is often discussed whether it is effective or not, even designers can be in conflict with the approach so in order to understand its strength, it has to be studied and practiced. There is a learning progression during the design thinking process and it affects the scale of effectiveness and quality of the results. Achievement and performance level can be related to the amount of time and practice helping to understand the limits of design thinking (Ericsson, Krampe and Tesch-Römer, 1993). Design thinking and the process models are neither recognized nor performed well in the majority of engineering industries and engineering schools. In an exam, it is not acceptable for a student to respond to a question by providing multiple possible concepts which cannot be measured from the true value point. On the contrary, students are encouraged to skip divergent way of thinking where there is not much of a place for creative alternatives (Dym et al., 2005). When ideation and research are not rich enough, the solutions are sought through the limited options. A regular designer has a potential to solve complex engineering or technical problems using design thinking and design research techniques (Katz, 2015). Design research therefore helps being a participant and observer in a contextual space. By observing the context and including different factors in the field, a designer can gain a first-hand understanding on different problems, thus gaining different perspectives regardless of the complexity (The Atlantic, 2011). It is important to understand the dynamics and the complex factors involved and therefore cannot be reduced to one method of data collection reliant on solely statistical analysis. As Wang, stated in her TED talk (2018, 00:08:17), “relying on big data alone increases the chance that we’ll miss something, while giving us this illusion that we already know everything.” Very few researchers suggest an increase in consumer understanding and trust in today’s ADAS and ADS technologies through human-centered design solutions and easily accessible information (Reimer, 2014; Fridman, 2018). However in these research papers, human-centered design was defined as targeting humans through interactive and personalized systems based on assumptions where there is no evidence of including the users in the design process. In human-centric design approach, root problems are identified through the insights which are gathered from people that are going to use the service or product. Removing the human factor and designing only based on the assumptions and self-thoughts is not a human-centered approach (Babich, 2018). Design thinking tackles the problems as system problems by taking all the factors into account as a whole. Different elements and factors such as hardware, software, procedures, people, policies, environment, ethics, etc, are all part of the problem which should be taken into consideration to create a solution on common ground (Owen, 2006).
When people have a conversation about the safety issues of SDC, the talk ultimately ends up with the popular question on ethics about how an SDC can decide whose life to take. The question is always built on an ethical dilemma; how can SDC decide whether to kill an innocent baby next to the roadside, or several people in the car if there is no way out. This popular ethical choice question is the new version of the well-known Trolley problem which was conceived by Philippa Foot in 1967 and was discussed by students in philosophy classes for decades (Lipson and Kurman, 2016, p.250). There can be multiple factors and conditions when talking about the trolley problem. Let’s consider for a moment that a car is controlled by a human driver and the driver faces a dog on the road. In this situation, the driver’s moral compass can make him or her do a sudden manoeuvring to prevent the collision but what if we change the conditions and the factors? If the scenario occurs on a very rainy day, would the driver take the same actions if there were children in the car? This is a never-ending discussion which is constructed on an ethical dilemma. AVs can make decisions based on the way they are programmed by the machine learning experts. If the algorithm output is programmed in a way to minimize the casualties, it will not observe the characteristic of people as part of the equation (Bonnefon, Shariff and Rahwan, 2016). AVs are also open to question from the point of product quality which determines efficiency and accuracy. Regulations may, therefore, need to be set on the quality of the components used in AVs, as well as how ethical issues are addressed in their design and algorithms. It is important to know how the decisions of AVs are justified and this can be possible through the transparency of AV companies (Holstein, 2017). Besides the popular discussions, the ethical experience design in an inclusive service to serve those with disabilities is one thing that AV companies should definitely pay attention to and carefully inspect. Without knowing how they really feel and think, an assumption based design may affect people in a negative way. The following findings are some of the concerns of blind and visually impaired people (Vymětal, 2013);

- Blind and visually impaired may need extensive and tailored information about security and evacuation procedures.
- Blind and visually impaired may perceive a strong need for safety precautions such as alarms and early warning systems for detecting impending dangers.
- Blind and visually impaired may feel more helpless, powerless, and dependent on others during and after a potential threatening event. Ignoring such facts when providing the AV service with the promise of inclusion is not ethical and these are the fundamental feelings and needs of people with disabilities to consider.
3 DISCOVERY
Curiosity and questioning have positively influenced every phase of this project from discovery to the execution of the idea. Designing things right is important and this can usually be achieved through the usability tests and feedback circulation following the problem definition. However, identifying the right problem and designing the right thing can be achieved through design research and early concepts and/or prototypes which is even more crucial (Tohidi et al., 2006). Asking the right questions can identify the right problems to tackle and the double diamond process model was therefore modified based on the needs and conditions throughout the research (Figure, 10). Considering time limitations, early concepts were executed in parallel to ongoing research where expert validation and feedback constantly developed an understanding of the existing AV issues as well as new opportunity areas. This allowed me to define one of the biggest issues in the AV industry through testing and failures of the preconcept designs. Considering time limitations, early concepts were executed in parallel to ongoing research where expert validation and feedback constantly developed an understanding of the existing AV issues as well as new opportunity areas. This allowed me to define one of the biggest issues in the AV industry through testing and failures of the preconcept designs.

Figure 10: Design thinking steps adapted from Design Council Illustration: Erdal, 2018
As previously mentioned, the goal of this industry research project was to increase the safety and public acceptance of AV technology as well as present solid evidence on how design thinking can play a major role in solving complex and technical problems. In order to reach these goals in four short months, intensive research was conducted in different countries and involved speaking to participants from different backgrounds. In order to comprehend the general knowledge about AVs beyond the popular literature as quickly as possible, primary research was conducted simultaneously to the secondary research. Two groups were involved in the interview process. The first group consisted of a diverse group of experts and members of the public and included engineers, roboticists, government and university officials and technical researchers, all interviewed at different phases of the project. The second group being AI systems and AI specialists. Working on several ideas and design concepts in the early phases helped the final research area and design question to be defined. Collaborating and meeting the experts as well as my mentor in person was a routine task for the final two months of the project.

Interviewees, Part 1

Ruben P.  Jens I.  Eetu I.
Jani K.  Rimma T.  Ilkay T.

The first group of the interviewees were organized to gather public insights and expert opinions on today’s AV technology.
The municipality of Amsterdam thoroughly conduct comprehensive studies of AVs and examine the social and economic aspects for future development. The government has designed future scenarios of AVs to be well prepared when the technology is commercialized while a team of experts constantly monitor unpredictable developments of AVs against all possible threats (Municipality of Amsterdam, n.d.). Polderman (2018), a project manager at the municipality of Amsterdam said, “We have an interest in this development and we want to articulate that interest. In the meantime, we observe the implications and early warning signals to take all the necessary actions because societal acceptance is extremely important for the existence of AVs and adequate and effective planning is therefore essential.”
In Denmark, in order to execute an SDC test, an independent assessor needs to approve a number of safety criteria within 6 areas including; general traffic, IT and data, data logging, vehicle technology, infrastructure and road technology, organization and resources, and risk management. It is then reviewed and approved by a task force consisting of: The Road Traffic Authority – the Police - The Attorney General - and the Road Directorate (Act amending the Road Traffic Act). Iversen (2018), smart cities innovation consultant at Denmark Technical University stated: “Government is concerned and strict about the AVs safety and even a small change like the door opening or closing speed of the autonomous vehicle should be changed upon the approval.” In Nordic countries, the authorities are precautionary about the execution of AV tests and unlike the U.S. test pilots, the AVs in these countries move very slow in a smaller area mostly with an operator staying in the vehicle ready to respond to any error or risk. Rutanen (2018), project manager of Helsinki Robobus confirmed, “We still have to have our own operator on the autonomous bus. We may test a remote operating system of the bus to deal with this issue.” Robobus, the autonomous public transportation project has been observed in person and its slow speed was not found attractive by its passengers who were interviewed. This fact was also confirmed by Chee (2018) during the ‘Drive Sweden’ conference which focused on self-driving shuttles in Sweden and provided the results from the trials, based on 6 months operation of transporting more than 20,000 passengers.
The participants were asked various questions in relation to safety and the trustworthiness of AVs, e.g. how they could feel in an AV which moves at a high speed.

"Going on the highway with the speed of 120 km/hour? I would have felt scared!"

(Kahkola, 2018), Infrastructure worker.

All the participants who were interviewed have expressed a common concern about the AV safety and the capability of the current technology. A contextual interview took place in an autonomous vehicle which was produced in 2018 and owned by a female participant. The participant demonstrated the self-parking feature both in a simple and complex environment; simple, being a wide space between vehicles and complex consisting of a more restricted space in which to park the vehicle. While the AV could easily park in a simple environment, the complex environment caused disengagement of the AV system and caused anxiety for the participant throughout the entire process (Figure, 11).
The participant stressed that she does not feel comfortable using the self-parking feature since she experienced some dangerous moments where she had to interfere to prevent the risk of hitting another vehicle or a pedestrian. As outlined in the literature review, once an individual questions their trust in a technology, they are more likely to avoid using it.

“It takes courage to let (autonomous) system to control the vehicle when your life is at stake”

(Tlekhurai, 2018), pilot.

“For a matter of time, I could keep an eye on the road when I’m in a self-driving vehicle. Once I’m used to it, could be easier.”

(Tunca, 2018) salesman and freelance translator.
There are various reasons for participants not to feel confident about AVs and some of these reasons were listed as; insecurity about the infrastructure, lack of trust on human drivers since they are unpredictable, distrusting of AI which operates the vehicle, joy of driving the car in a conventional way. During the research in Copenhagen, passengers on an autonomous train were observed on the front side of the vehicle where the operator is located in a conventional train. There was no sign of a discomfort or fear even though the AV was as fast as any conventional train.

Common insights of the interviews and observations, Part 1

- If AI and people can work together, that is the most reliable option.
- Supervised or private AV roads with a smooth infrastructure make participants feel safer.
- Pilot projects such as autonomous shuttles gave people the impression that all AVs are slow and not advanced enough.
- All participants were concerned with the idea of professional drivers losing their jobs and this ethical issue should be taken into account.
- People have very little knowledge on AVs other than from what they learn through the media and this can be considerably misleading, thus not fully highlighting AVs potential.
Following the interviews which were conducted with the first group of participants, a short survey regarding the previous findings and insights was designed and shared with people. Of the study population, one hundred and twenty people completed the survey and the results were mostly consistent. Majority of those surveyed indicated that AI and human coexistence has the best outcome (Figure, 12).

**AV & AI, survey findings - 120 responses**

**Let’s start with your age?**

- 18-28: 65.6%
- 29-39: 26.3%
- 40-55: 3%
- 56-65: 1.7%
- 66+: 0%

**Do you think autonomous vehicles are safe?**

- Yes: 30.6%
- No: 46.7%
- They drive better than my grandma: 3.3%
- Depends on the technology provider (brand/company): 16.7%

**Would you give up driving a car if autonomous vehicles were 100% safe?**

- Yes: 39%
- No: 61%

**Do you think AI and human coexistence has the best outcome?**

- Yes: 35.4%
- No: 11.8%
- It might be risky in the future: 25.5%
- Maybe: 27.2%
3.2 Opportunity and ideas

Enhancing the road infrastructure using constant monitoring and ‘infrastructure to vehicles’ (I2V) technologies:

This could create an opportunity to observe threats on a wider area since the devices would be mounted on higher levels similar to road lights. While this is feasible in today’s technology, not all AVs use C-V2X features and as previously stated, it will not be profitable enough unless the majority of the automotive companies and vehicles in traffic switch to the C-V2X technology. In order to solve this problem, an alternative communication model which works through the traffic signal like a type of sign language can be designed. A simple and universal AI language can be a communication alternative to inform AVs, and more details on this will be given along with a concept and can be found in the recommendations section.

Designing an inclusive service which is profitable for elderly and blind people:

This opportunity area has been explored mostly through secondary research and assumptions based concept design (Please refer to Appendix 1). In order to develop a realistic and human-centered solution, and to test the concepts, it was necessary to spend a very long time to communicate and collaborate with people who have disabilities and mobility difficulties. Due to the time frame and ethical concerns, this idea was not considered to move forward with.

Enabling the coexistence of AI and Humans in AVs:

Collaboration of AI and humans demonstrate a significant impact in different industries. As an example, Harvard’s medical research team developed a method to diagnose cancerous cells using the AI which was proved to be 92% accurate in comparison to a pathologist who is classed as having 96% accuracy. Even though humans can make more accurate diagnoses, the collaboration of AI and humans raised the figure to 99.5% accuracy (HMS Harvard, 2016). Exploring the opportunity areas within the coexistence of AI and humans was considered to be the best option to move forward with. However an additional research and interviewing AV/AI experts considered to be necessary to understand the context and technical aspect so as to tackle real engineering problems.
3.3 Understanding the AI
As previously outlined in the methodology section, different AI systems both in physical and virtual form have been interviewed and tested in order to understand the AI including its limits, patterns and the way it works. Two of the participants were humans with the expertise of machine learning and AI and were asked to adopt the character and personality of an AI and answer questions from the AI’s perspective. Among real AI systems, Mitsuku, Google Assistant and Furhat were the most advanced models to explore. The interview questions were designed to understand the complexity of an algorithm, learning capabilities and perception of the AI systems.

Participants stressed the importance of being fed with diverse data in order to develop their intelligence and perception.
Mitsuku and Google assistant can learn and remember things they are taught. During the test interviews with Google assistant, information such as my favourite food, the name of my best friend and my birthplace were taught easily.

“I need **diverse data** to grow. Diverse data is like food to me. You need all sort of nutrition”


“There are many scenarios that cannot be solved with my existing data”


“I am definitely smarter than a toaster. They do really shine when you are in the mood for toast, but I am a lot more versatile”

Google Assistant/AI - 4.

“I love to learn. I am trying to learn all about you humans and your behaviour”

Mitsuku chatbot/AI - 2
Google Assistant and different type of chatbots, specifically Mitsuku were engaged in hour-long conversations. A social robot, Furhat and many other AI systems are able to contextualise and therefore communicate in a certain area of expertise with most having a supervised algorithm.

- The different types of AI agents all require constant learning in order to improve performance and accuracy. Training data is, therefore, the most vital element for an AI which should be diverse and enriched to increase the ability of perception.
- AI can be personalized since it has the ability to learn from people.

3.4 Problem definition

The issues within the AV field became more clear following the studies and the first two sets of interviews of experts as well as participants from the public. Some of the core insights gained from the entire research follow as:

- People are not informed good enough about AV technologies and often misled by mass media.
- AI agents require constant learning in the AV industry and this job is being done through a massive amount of people spending too much time and money every day.
- Technology-based modifications of infrastructure have a great potential to reduce the accidents even without the network connection. Well designed and quality infrastructures also affect people’s thoughts on AV use.
- Different AV companies and technology providers have different success rates which also determine AVs’ reliability based on the brands.
Design question

Taking the AI, consumers and AV companies into consideration, the design question had to match all stakeholders’ needs through a binding experience. For each stakeholder a separate behavioural persona was created to help document the needs, feelings, and motivations and based on a binding empathy map the design question has been formed;

‘how might we design a consistent feedback circulation between AI and people?’

‘Five whys’ method by repeatedly asking the question ‘why?’, ‘why this is important?’ was used to evaluate and validate the symptoms that lead to the root cause of the problem.

3.5 Ideation

The ideation method; Crazy 8’s was used for fast sketching where eight different ideas in eight minutes should be created based on the design question. The exercise was very useful since it helped to silence the inner critic and gave more space for creative and diverse ideas within a limited time. Two of the ideas were then chosen to create more detailed sketches on separate papers.

Idea 1

The first concept targeted the consumers who have SDC. Basically, a bracelet which can constantly monitor the heart rate during the journey would detect the change in heart rate and report critical differences to the data science team. This report could include, vehicle speed, a visual from the environment, and the data, based on the perception of AI/computer vision.

How would this help? This could help to analyze by only the moments when an AV takes a risky action like disengaging off the route or failing to recognise an object and it can help to reduce workload. AV can also track the condition of the passenger through the heart rate monitor and modify the speed and the way it drives accordingly.
Idea 2

The second concept targeted the passengers who travel on autonomous shuttles. The idea is to help AI to learn through a gamified platform where passengers can validate the object detection results. A platform where passengers can basically connect to live video stream showing an AV’s surrounding in real time and validate the objects which AV encounters (Figure, 12).

How would this help? This could constantly advance the accuracy rate of the AI on the same track and in different conditions. The passengers would have been informed about different facts of the AV through the game and get a travel discount for using the service which could be quid pro quo for all stakeholders.

Feedback and building on the ideas as a team

The two ideas were considered to be presented as ‘sacrificial concepts’ to participants in which they would choose which one they liked the most even though the first idea was personally found ideal to work on to develop further. ‘Sacrificial concepts’ is a design thinking method which was found by IDEO, a global design firm. In order to use this method effectively and gain valuable insights, two or more early-stage concepts should be presented to the users and egos should be left behind. These concepts are a great way of sparking a discussion around the presented ideas and work as an idea generation mechanism which is a great way of improving the concept by building on each other’s ideas (DEI613, 2016).
“The issues we face are so big and so challenging that we cannot do it alone, so there is a certain humility and a recognition that we need to invite other people in.”

— Paul Polman, CEO, Unilever
Two participants were selected based on their sufficient knowledge in AI to be able to discuss technical details and build on each other’s ideas. Sacrificial concepts were introduced, and the discussion method was explained. Each participant had around fifteen minutes to think and note their feedback on sticky notes for each concept. The discussion started after the feedback was written for each concept.
The concepts were discussed based on their feasibilities, desirabilities and viabilities. After spending nearly an hour on brainstorming and building on each other’s ideas, a final concept was created. The first concept which was the heart rate monitoring idea was eliminated and the second concept with the gamification idea has been improved.
AI training platform can be completely gamified and published as a real game which can be accessed globally. Anyone with smartphone and tablets would be able to play the game and win real prizes such as a driving discount for ‘Lyft’. Additionally, games can also be used by industries like defence, education, scientific exploration, health care, emergency management, city planning, engineering, religion, and politics apart from leisure. They are called serious games and their main purpose is to train, investigate, or advertise (Ioana Muntean, 2011).

“"I would definitely play the game especially if it gives me a public transportation discount”"

(Klaesson, 2018) Roboticist

How does the training platform work? Users would be able to watch images or video footage on their phone and the goal is simply to click on the objects which appear with a question mark tagged on them, indicating that a validation is needed. The game can be designed in a fun and simple way which makes it easy to play where users can both teach the AI and also be taught about different technologies including AV based facts. As confirmed by Russell et al., (2008) “once there are enough annotations of a particular object class, one could train an algorithm to assist with the labelling. The algorithm would detect and segment additional instances in new images. Now, the user task would be to validate the detection”. The feasibility of the idea was confirmed through the secondary research and expert validation.
In order to have a collection with a big amount of visuals annotated with maximum efficiency and quality, the annotation tools need to be made operable easily, the annotation progress needs to be monitorable, and concurrent annotation sessions need to be supported (Yu et al., 2018). One of the most efficient and cost-effective methods for image annotation in high quality is to outsource the data training jobs in poor countries. Samasource pays $9 per day for image annotation jobs and have nearly one thousand employees in Kenya (Figure, 13) which provide trained data to companies including Google, Microsoft, Salesforce and Yahoo (BBC News, 2018).

More than a million users can be encouraged to register and play the game by spending a very small amount of what companies pay for image annotation today. Having one hundred thousand active users means doing the same job tens of times faster and that much cheaper. In one example, Samasource’s one thousand employees can continue their jobs as supervisor to check the suspicious inputs or to validate the annotation of unusual and rare objects. Even doing so, the company can analyze a hundred times more data with almost the same cost and more important can finish a year-long data training job in only three days. By implementing this idea all three stakeholders’ needs can be addressed where AI can constantly learn, people can be informed about the AV facts and adopt the technology easier, and technology providers spend less money by being able to advance AV algorithms faster than ever. The solution has the potential to connect all the stakeholders through a binding experience.
Due to time limitations, the first prototype was crafted using transparency films where each one of the films was used as a layer of an interactive screen model to imitate action steps (Please refer to Appendix 2). All three participants who have tested the prototype expressed their expectations and found the design easy to perceive.

Several questions were asked about the game mechanism by the participants and those questions along with the ideas were embraced to design the final prototype in digital format. Final concepts and low fidelity prototypes have been validated by the experts and advisors and the digital prototype was built using the InVision software. In the last week of collaboration with ITRL, the idea and methodology were pitched to the staff, PhD and Master’s students as well as senior researchers who work on AVs at the Royal Institute of Technology. All the feedbacks and the security-related questions were noted for future developments following the delivery of this research project.
As noted earlier, ethics play a major role in research, especially research that involves engaging and interacting with humans. In order to proceed with ethical considerations in mind, participants needed to be fully informed of the procedures and processes involved in the study, as well as understanding of the main purpose of the project. The purpose being to train users with autonomous and related technologies, supported with facts obtained through scientific research. It was stated clearly that the most important purpose is to train an AI algorithm which can be used in different industries. The future product is planned to be developed in a way that training can be optional and the user may have the right to withdraw at any time. The future aim of the data collection process is planned to provide a fun and educational experience both for users and AI where engagement with the game is always at the participant’s own discretion. Publishing any future application related to the game is planned to promise the transparency and protect all the user rights without any misdirection and misuse of user’s information.

“It is simply not acceptable to take the attitude that, ‘I just make the tools. I can’t be responsible for how they are used.’ The reality is, the design can have a huge impact on channelling usage along certain paths. Those of us who design such things need to make the best efforts to make sure that those paths conform to our ethical compass.”

Bill Buxton, principal researcher at Microsoft Research

Because the design of the digital applications can easily mislead users, designers are responsible for the experience and therefore play a crucial role in maintaining ethical considerations of both user, the experience, and the design itself, as the design will affect the way the user perceives the experience (Jensen and Vistisen, 2013). The implications of the game application were discussed with the experts and colleagues. Implementing the solution successfully might increase the data training speed dramatically and therefore enable AVs to be commercialized faster than expected which results in professional drivers to lose their jobs earlier than expected. However new opportunities for low qualified workers may rise in automotive-related areas. A possible profession area with existing research applications is expected to be ‘remote vehicle operation’ since companies will need to establish vehicle control centers, to monitor fleets of AVs and take over the control when a vehicle has a malfunction or needs human support. Working on new technologies might be considered ethically wrong since they remove some traditional work from the market, however, they always created new occupations in our history (Boston, 2017).
Following the four months of rigorous research, significant findings emerged, and it became apparent that two main factors play a vital role in the technological advancement of artificial intelligence. Furthermore, the findings show a positive aspect regarding public perception and potential for human contentment living alongside AI systems providing the design thinking can ensure a safe, reliable and efficient programme for AVs and AI. Technologically advanced road infrastructure and diverse datasets which affect decision-making processes, particularly the decision-making algorithm, appear to be the way forward to improve and develop the AI to meet all the needs that are required to fit into society, safely and with acceptance from the general public. Diverse and enriched datasets are the most important issues that require improvement, especially the dataset issue, however, by adopting the suggestions set out in this research, could potentially lead to a modern day technological revolution. All the existing solutions on image annotation target the technical improvements whereas this proposal offers a change in methodology by providing a gamified platform to crowdsourced the workload in a fun and educational way. Implementing such a platform successfully will increase the trained data input speed dramatically which means there would be a cost reduction whilst significantly enhancing the efficiency of technology development. However, the execution of the idea requires intensive work and collaboration with data scientists, machine learning engineers, computer vision and AR experts, digital experience designers, game developers, visual artists, etc. This project also aimed to tackle engineering problems in order to increase the safety of AVs. This was evident through the use of design thinking approach which focused on a desired, feasible, and viable service because, as outlined before, technology and the general public are not entirely ready for SDVs. Mistrust in technology can result in a reluctance to use AVs or their ADS features and may slow down further developments. During the contextual and in-depth interviews participants expressed their thoughts and feelings on AI and AVs, emphasising the most important aspect of their uncertainty as the lack of technological advancement and conditions of infrastructure, which was deemed much more of a concern than how nicely a product is designed or advertised. Such concerns should, therefore, become a priority and addressed immediately with optimum efficiency. Despite being the central and most fundamental aspect of a desired future with commercialized AVs, humans cannot be treated as test tools. They will not risk their lives without an optimal level of advancement and in technology with guaranteed safety. Human needs, feelings and thoughts should, therefore, be considered in every action step in order for it to be considered effective progression. Moreover, data security ethics, AV decision making processes, and consequences of a driverless future should also be taken into account by governments and technology providers. AI can, therefore, be seen as having the potential to not only change people’s perceptions but to change society as a whole as advancing technology becomes more influential and highly regarded across the globe. This research topic is reserved for future work and will be further developed in order to advance knowledge and understanding in this subject.
As previously outlined in the ‘opportunities and ideas’ section, infrastructure has a vital role in increasing the safety and reliability of AVs regardless of how advanced AV algorithms can get. Some companies such as VZ investments focus on physical modification of the road infrastructure, creating a safe area since 50% of serious collisions, and 20% of fatalities occur in intersections. An intelligent intersection functions as a ‘modified intersection’ using I2V communication and can reduce collisions (Grembek et al., 2018). Current research focuses on using the I2V technologies by mounting the sensor on to the infrastructure and monitoring the surroundings which then informs the vehicles. However, this solution will only work for the vehicles that have C-V2X technologies. A simple and universal AI language in the form of traffic signal light can be a communication alternative to inform AVs both with and without C-V2X features. Infrastructures and AVs together can communicate with each other through the psychophysical laws as seen across a spectrum of mammals, fish, birds and insects which show similar dynamics to human decision making and communication also referred to as Pieron’s and Hick’s law (Reina et al., 2018). The infrastructure can also observe blind spots and inform all the agents of potential threats using different numbers and combinations of lights. Infrastructure can further increase the perceptive abilities by scanning threats in a wider area. It then uses this information to evaluate the conditions through the data exchange between neighbour infrastructures and the vehicles nearby, therefore improving the decision-making abilities.
BIBLIOGRAPHY


WIRED (2018) *Inside the Races That Jump-Started the Self-Driving Car*. Available at: 

Last Century, the Present Scenario and the Expected Future of Autonomous Vehicle Technology. In: *In
Proceedings of the 12th International Conference on Informatics in Control, Automation and Robotics

The Economist (2018) *Autonomous-vehicle technology is advancing ever faster*. Available at:
https://www.economist.com/special-report/2018/03/01/autonomous-vehicle-technology-is-advancing-ever-fa-

Engineering (2018) *The Road to Driverless Cars: 1925 - 2025*. Available at:

Available at: https://www.techemergence.com/self-driving-car-timeline-themselves-top-11-automakers


sites.ieee.org (2018) *You won’t need a driver’s license by 2040*. Available at:


Entrepreneur (2018) *3.5 Million Jobs Are on the Verge of Extinction. Here’s How Entrepreneurs Can Save

NBC News (2018) *Millions of jobs are on the line when autonomous cars take over*. Available at:

Ottawa: Standing Senate Committee on Transport and Communications, p.34. Available at:


Arbor, University of Washington, University of California Berkeley, Samsung Research America, and Stony Brook University, p.1.


*Act amending the Road Traffic Act. Law No. L 120.*


**VISUAL REFERENCES**

Freepik (2018) Autonomous vehicle illustration adopted from Freepik. [image]


Erdal, Y., (2018) Idea (Sacrificial concept) 2. [image]
Appendices 1

Concept Design for blind people
Appendices 2

Transparency films
Appendices 3

Early concept

Chatting with one of the AI chatbots